

# An Inertial Tale

---

## Open access

In principle, most researchers would subscribe to the view that the outcomes of all publicly-funded research should be available to anyone in the world, without charge, other things being equal. This is the principle of *open access* (or *OA*) – a code term for this state of affairs as applied to a particular outcome. Let's take this attitude as a given, and not argue it. I am fully aware that there can be reasons for research outputs to be kept confidential, but these comprise a small number of exceptions. Generally, researchers want to share the results of research with others, and take their benefits in reputation and perhaps status.

While open access prevailed in the very early days of science because the community of scholars was small, the ideal of open access was not achievable, or achievable only on a limited scale, for almost all of the subsequent history of science and scholarship. The costs of copying and dissemination simply prevented research outputs being made available to all, free. As a partial substitute, a publication industry grew up around the research community. For my purposes, this can be characterized as the scholarly journal industry, while recognising that the proceedings of some conferences are also part of the industry. Universities and other research organizations basically fund this industry, through their subscriptions to journals or journal bundles. Access to research outputs is restricted to those with the capacity to pay. Researchers provide their research outputs free to the industry, and they provide free labor in quality control (peer review). They expect no direct monetary reward.

However, thirty or so years ago, the ideal of open access took a great conceptual stride forward. Computers had already made the copying of any research article very small – so small as to be less than the smallest coins of currency (cents and pennies), and basically negligible. The costs of storage of the article dropped dramatically too, and while initially higher it is now also negligible. For example a 1 TB disk might cost \$100, meaning that a large 10 MB article requires storage costing 0.1¢! A physical space the size of a book can hold 100,000 articles of this size, so physical space accounting is equally small. The final piece of the puzzle fell into place with the worldwide network (the Internet, or net of nets) which made the incremental cost of dissemination similar. Again, indicatively \$50 for 500 GB download (both per month) means a large article of 10MB costs maybe 10¢ to access. In Australia, it is likely that 10c will be the smallest coin minted quite soon, since the metal value of a 5¢ piece is approaching 5¢ and will eventually exceed it.

## The two roads

Despite the time that open access to all publicly funded research has been possible (30+ years), it is well-known that open access has not yet become universal, nor even are the majority of research outputs openly accessible. Several people have devoted significant parts of their career and effort to trying to make this happen. In this essay I look at these two roads and examine why they have failed to attract enough travellers.

Conventionally the two roads are called the *Green Road* and the *Gold Road* to open access. I don't know how the names were chosen but many readers take a subtle implication in the naming that the Gold Road is better, though this is not universally agreed. Let's examine each road in turn.

## **The Gold Road – Open Access Journals**

Traditional journals are funded through subscriptions, and they make their articles available only to subscribers. Frequently the articles are available online to subscribers, and in many libraries at present the paper copy of the journal is regarded as a nuisance and sometimes even tossed into a waste bin. It costs more to catalog, shelve and store than it is worth. Typically, nobody ever consults it. It is perhaps an insurance against the publisher becoming bankrupt.

The Gold Road to open access describes a situation where a journal provides open access via the Internet, from the day of publication and for perpetuity. In other words the open access is provided by the journal publishing industry and is as permanent (or not) as the journal. The journal may be a transformation of an old one into a new business model, or a new activity like the PLoS stable of OA journals in life sciences. Of course, while the publisher can take advantage of the low reproduction and dissemination cost, it would not be much of an industry if it did not have some real costs. Largely these would be in editorial and certification (peer review) activity, though there would also be some minor costs in archiving and ICT. OA journals should have significantly lower costs than journals which appear in print and cater to subscribers.

Typically, university librarians and research managers like the concept of OA journals (the Gold Road) far better than the Green Road. After all it preserves the idea of a journal, even if the physical representation has withdrawn into cyberspace. With the idea usually comes the usual baggage of regular issues, ISSNs, subject identification, contents listings, citations and citation metrics. Actually most of these are paper artefacts hanging on despite the change of circumstance, but still they are comforting to librarians and researchers. Their continued existence is itself a hint that electronic dissemination through the Internet has not yet matured, and made an adaptation to the medium.

However OA journals are not a solution to the problem. They are slowly growing in number, but remain a minority of scholarly journals. If OA journals are going to bring open access to the world, we had better not hold our breath, because it will take a long time. A transformational switch is highly improbable, therefore. Why? Because the business model for an OA journal is far from stable, and the source of funds is not clear. As long as most readers' funds (actually the readers' libraries) are locked up in subscriptions, they are not available for paying author-side fees. The articles available by Gold OA journals are not the same ones as available by subscription. Largesse from funders in the form of extra Gold Road subsidy payments only change the situation slightly by levelling the playing field for some authors. They do not create an imperative reason for change. The key issues here are:

1. Gold Road OA journals are generally not the top-ranking journals in the field, with the possible exception of the PLoS biology stable, and journals about the Internet.
2. The domain of toll-access journals and open access journals generally overlap, and a research organization needs to have read access to both. While access to the Gold journals is assured, the subscriptions to the toll-access journals (or bundles thereof) simply must be maintained.
3. On the converse side, author access to being published in toll-access journals is assured (publishing is provided free), but access to publishing in the OA journals often costs funds, and there is nowhere for this to come from.

## Hybrid practices

A number of publishers have adopted hybrid practices (in other words OA and non-OA). These are not significant changes in policy, nor transitional to a fully OA production. They pretend to pander to the OA community by challenging it to pay for access. The subscription (reader-side) fee still applies, but by supplying an extra payment (author-side) an article can be made immediately OA.

There is actually no limit to what can be described as hybrid OA. Many publishers make articles OA after an embargo time (a hybrid practice), while others do not. Other publishers choose a random few articles to be made OA immediately (as a teaser); my description as ‘random’ also includes selecting tantalizing articles. Even Amazon.com adopts hybrid practices with respect to books, often making selected small parts open access, while of course the abstract of a journal article is nearly always made open access. All hybrid practices can be seen as either transitional or as a sop to the OA lobby, while keeping the lucrative subscription monopoly running as long as possible. No more time need be spent on hybrid practices. They are not fundamental nor transitional, but rather indicative of change. I will waste no more space on hybrid practices.

## The Green Road – Repository deposit

The main alternative route to OA is the Green Road – authors deposit their articles into an Internet-connected repository, from which they can be retrieved. There is no mention of publishers in the basic idea; deservedly since authors provide publishers with the raw material to make profit, free of charge. The Green Road has been pushed by many people, but is most closely associated with Stevan Harnad.

The idea has undergone some transformation over the twenty-or-so years it has been in existence, but basically it goes something like this:

1. Universities and research institutions (eg CERN, CSIRO) establish and operate repositories on the Internet.
2. Authors upload all their journal publications while in the employ of the institution to the repository. It is important that the articles be journal-certified, even if the journal is traditionally paper-based.
3. The repository is connected to the Internet, and all content on it is publicly indexable by robots, and downloaded by interested persons.
4. Gateway websites harvest summary data from the repositories (through special interfaces) and provide search facilities for the articles, worldwide.

This may be regarded as a basic strategy. There are many things that can be said about this model and many twists that can be applied to it. For example, what version of the article is deposited? Who pays for the running costs of the repository? Who can deposit? Is the deposit subject to librarian oversight and change? Is the institution a publisher thereby? What rights over this process do journal publishers have?

At first sight, this model looked so obvious and desirable that it was simply promoted in this form. Rapidly, it became obvious that it was not being taken up and three key innovations were added to the model.

Firstly there was the ID/OA model, becoming used as a mantra for promoters of the Green Road. The initials stand for *Immediate Deposit, Open Access as soon as possible*. The basic idea is to recognize that all articles are written on a computer at this time, and the immediate deposit captures the born-digital file at the earliest possible time. The object captured is the author's

accepted manuscript, and the action is taken as soon as the article is accepted, before it can be lost. The second part of the mantra simply says to make the article open access if that is possible; otherwise it is still captured as closed access. Readers of the repository cannot access closed access articles, though the author usually can access his or her own article. The ID/OA strategy ensures that the articles are available after the passage of an embargo (entered at the same time and automatically actioned), though as a side-effect this provides the author a convenient zero-weight library of his/her own articles. This actually proved to be a big advantage for highly active researchers.

Secondly, repositories were fitted with a 'Request-a-copy' button, whereby a reader could initiate a request to the author for a reprint copy of a closed access article, with just a click, the entry of an email address and a second click. The author could approve the request by a single click on a generated email. This innovation mainly ensured that captured even closed access articles could be usefully used, under general conventions regarding the provision of reprints by authors.

Thirdly, a campaign was commenced to make ID/OA deposit in the institutional repository *mandatory* (in other words, required). It had become clear that when deposit was left voluntary, the deposit rate stayed at about 10–15%, even if the repository managers provided a great deal of encouragement and persuasive techniques. These in fact made little difference, and the deposit rate stayed at this low rate or a few percentage points higher regardless. It seemed that persuaded researchers failed to deposit after a while (backslid) and that as fast as people could be persuaded, others dropped by the wayside.

The questions that have to be answered in respect of mandates were (and are) never-ending. The mandatory element was only reluctantly adopted, and it remains a main bone of contention laying across the Green Road. Why is compulsion necessary? Isn't compulsion repulsive? How long will it be necessary? Doesn't the mandate negate the Green Road completely and make the Gold Road the only avenue a free researcher could possibly support? If a mandate is so important, why have the administrators of research institutions not adopted it universally? Does the mandate require exemptions, say if the publisher does not approve a deposit? What is the difference between a funder mandate and an institutional mandate? Why is there no example of a subject-based mandate? Where does copyright fit in the mandatory regime? What is the validity of the agreements authors sign with publishers, without reference to their employers? What carrots or sticks do repository managers use to enforce the mandate (if they do)? And so on...

## ***Confession***

I admit myself to having been persuaded by this argument and having lent my effort to trying to make it work. I analysed an early adopter of a mandate (Queensland University of Technology, with Tom Cochrane's policy) and showed that the mandate made a difference. Deposit rates rose to 80% after 2 or 3 years and stayed there. This experience was replicated elsewhere. Mandates worked at QUT. I remain convinced that mandatory deposits, where achieved, should continue, and should be encouraged. However, I have now equally convinced myself that they are not the answer to making the world's research open access.

The one place where mandates are 100% effective is in PhD theses and dissertations. Candidates are used to being told what hoops they must jump through in order to graduate, and even worse, they have no choice. The change to PhD deposit regulations turned out to be truly minor in the University of Tasmania, and coincidentally cheaper for the candidate: a successful graduate has to deposit one electronic copy of the thesis with the Graduate Research Office, rather

than three hard-bound paper copies. In all cases the copy or copies is transferred to the University Library – an electronic copy is put straight into the online repository. The regulation change also opened the possibility of a thesis being irreducible to paper, perhaps containing animations, video or hyperlinked parts. Compliance seems assured, now and in the future.

## ***Inertia***

In analysing the situation in all its complexity, I was reminded of the *Foundation* trilogy and its 7+ follow-ups, by Isaac Asimov<sup>1</sup> and subsequently others. For those not familiar with the works, Asimov originally wrote three Science Fiction novels based on the following assumptions, the first two of which are common sci-fi assumptions to make the stories possible:

- Many (or most) habitable planets in our Galaxy (the Milky Way) are populated by sentient beings.
- Traffic between habitable planetary systems is possible by ‘jumps’ through ‘hyperspace’ which instantaneously move a vessel from one point to another. Long trips involve multiple ‘jumps’. There are other restrictions of no interest to us here.
- Our Galaxy is populated by a single sentient biological species (humans).
- The Galaxy is initially controlled by the Galactic Empire from the planet Trantor, and the Empire has commenced a Decline and Fall, though the signs are not recognized.

A mathematician, Hari Seldon, recognizes the collapse and develops the science of *psychohistory*, which enables prediction of the behaviour of large masses of humans. As the Galaxy contains probably several Tera-humans, prediction is presumably pretty good. However, one of the requirements is that the population be unaware of the predictions, as this knowledge would affect its behavior. The science predicts 10,000 years of misery for the Galaxy after the Fall before a new stable system of government is established. Dr Seldon sets out to reduce this period. In his eventual trial for sedition, Hari affirms that he cannot stop the Fall of the Empire, and most anything he or anyone else can do will merely accelerate it. But he can shorten the period of chaos from 10,000 years to a mere 1000 years, by making a few key changes to galactic society.

The key issue here is that the Fall of the Empire is already in motion, with a dreadful inertia. Nothing can stop it; most things speed it up. We might compare this with the motion of a ship being launched on the ways. Once it is in motion, stopping it is very difficult, but creating more mayhem is easy. The inertia of a ship is a big thing. Yet, what happens afterwards can be affected by small things, given enough time to take effect. The questions are how much time we have, and how much effort should be applied. Inertia takes its place here too.

And so to the Green Road. The inertia militating against change is comprised of the hugely dominant prevailing view (by researchers) of journals as the only way of publishing research results; of paper printing as the only (or prime) archival record; of kudos deriving from being published in key journals; and of a known effort and money to be applied (writing an article, submitting it, addressing criticisms; and copy-editing of galley proofs). The simple inertia of the social system has proved to be too large for the open access view to take root and prevail. Never mind that the Green Road requires perhaps 5 or 10 minutes extra work per article in an activity that spans months if not years. Many millions of researchers over tens of thousands of institutions know that journals are their way to get published; they never see the costs; and they have little interest in the publishing

---

<sup>1</sup> *Foundation, Foundation and Empire & Second Foundation*, Isaac Asimov, 1948-52.

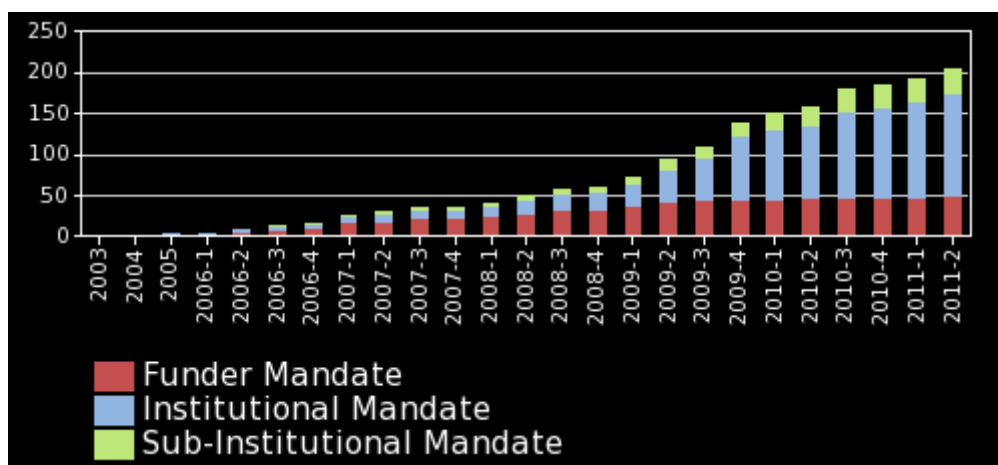
industry tail attached to the research dog. They are quite happy to leave that to the publishers and paying the costs to the librarians.

With this perspective, I see the history of the Green Road in a different light. In its original instantiation, the target was to convince researchers to deposit their research and to convince institutions to establish and pay for the (small) running costs of an institutional repository. The second half of this proposition achieved moderate success, but the first half was an abject failure. The inertia of millions of researchers all over the world was simply too great for the efforts of the relatively few open access persuaders. As soon as a persuader left a converted researcher to start on another, the first researcher started to backslide, with the example of millions of unconverted researchers around him or her who did not spend time on the deposit chore and seemed no worse off. The net result was that the participation rate did not increase even linearly, but reached a stable plateau where recruited researchers balanced those leaving, maybe at about the 10% level at most.

As this was realized, attention turned to convincing research leaders, in the hope that since they comprise a smaller subset, they might be more easily induced to maintain participation, and became ‘champions’ leading to auto-conversion of their fellow and junior researchers. However, once again, participation stabilized at a plateau level, far below that needed to reach a tipping point. The message is simple: the inertia of the researcher: publisher system is too large to affect by direct action such as this. More economy is needed, because OA proponents are very thin on the ground.

The next response was to go even higher in the researcher hierarchy to the research managers, and to try to get them to mandate deposit, in the interest of the institution. There are perhaps 10,000 universities to be tackled. Mandation seemed to work where it was applied, but few mandates were applied by truly convinced managers (QUT is a key exception to the rule). The mandates abound with exceptions (‘deposit as long as your publisher allows you’) and surreptitious wishes for other solutions (‘the Gold Road is the only long term answer’), besides having problems with compulsion (‘what sanctions can I apply to a researcher who ignores or forgets the mandate?’) and copyright (‘what legal problems might the institution encounter if it deliberately ignores copyright?’ and ‘would I be personally responsible in that case?’). Australia virtually only has one clear example of a mandate in 34 institutions, and yet most institutions have repositories, almost all practically empty.

Perhaps a very clear example can be found in a recent email<sup>2</sup> by Stevan Harnad, where he included the following graph probably from Alma Swan:



<sup>2</sup> Private communication, June 2011.

Stevan points to a recent spurt in institutional mandation in 2009-2010. I see the graph as indicating yet again the inertia of the research system, tending to level out on a plateau. Funder mandates have ceased to grow, and institutional mandates are slowly ceasing to grow. As the effectiveness of past mandates is not examined, the situation may be even less rosy. I know quite well that even the limited sub-institutional computer science mandate at the University of Tasmania has now become forgotten and no longer honoured. The database still shows ancient history.

## **In Conclusion**

It has always been quite clear that the Internet offers the possibility of returning to the beginnings of science, before journals were invented. Research results were available to all for a minor cost. Open access offers this possibility to the vastly expanded set of researchers, worldwide, and for a pittance.

However it seems equally clear that the efforts of the 1990s and 2000s have failed to bring this desirable state of affairs into existence, and the rate of improvement (if it exists at all) does not seem likely to result in open access in my lifetime. The publishers can rest easy, not even worrying about being the tail wagging the research dog, but knowing that the inertia of researchers and research managers guarantees the continuance of toll-access journals into the indefinite future. The only minor worry on the horizon is that of Gold Road journals, and they show no sign of growing at anything other than a slow linearity, when research outputs are increasing at a much faster rate. The situation is not much different from that of climate change activists, except that there are vastly more of them than there are of open access activists.

My conclusion is that I need to rethink the strategy of open access, so as to apply the minimal changes that will set in motion an unstoppable set of changes, and to do this knowing that the resources of the open access movement are very limited. Hari Seldon would have approved. In the next essay, I turn to examining the nature of scientific revolutions, and to try to bring out what characterizes them, and if this knowledge can provide the lever we need to achieve open access in a short time.

